The NASA P-3B aircraft was equipped with in situ probes during the NASA ObseRvations of Aerosols above CLouds and their intEractionS (ORACLES) field campaign. The probes included a Cloud and Aerosol Spectrometer (CAS), Cloud Droplet Probes (CDP), a Phase Doppler Interferometer (PDI, serial number 0491), a 2-Dimensional Stereo Probe (2D-S, serial number 012), and a King hot-wire probe (model KLWC-5, serial number SN-PMI-1058-0704-86). The CAS was a part of the Cloud, Aerosol, and Precipitation Spectrometer (CAPS, model AAA-0009, serial number 5). The P-3 carried a single CDP (CDP-A, serial number 0901-48) during the 2016 Intensive Observation Period (IOP). A second CDP (CDP-B) was added to the P-3 for the 2017 and 2018 (serial number 1206-070) IOPs. CDP-A was replaced by a different CDP (CDP-C, serial number 0604-006) for the 2018 IOP.

The probes were calibrated by the manufacturers before and after each ORACLES IOP. Instrument performance was monitored during the IOPs using calibration tests and auxiliary data, such as temperature and sensor voltages, were monitored during the research flights. Flight legs through aerosol plumes with high (greater than 1000 cm^{-3}) aerosol concentration (N_a) were conducted during ORACLES. These plumes contained soot particles that could adversely affect the quality of measurements, especially for the 2D-S. This was addressed by cleaning the optical lenses of the probes with isopropyl before each flight. Gupta et al. (2021) examined the 2D-S measurements and discussed the data processing techniques used to identify and remove data artefacts.

The objective of this supplement was to compare data sets created using measurements from different cloud probes used during ORACLES. The focus was on droplets with diameter (D) between 3 and 50 μm since the CAS, CDP, and PDI measured droplets over this size range. The

differences between droplet concentration (N_c) and liquid water content (*LWC*) from the CAS, CDP, and PDI data sets were determined. While they may, or may not, be within the uncertainties (Baumgardner et al., 2017), the differences between the data sets were quantified to illustrate that using one instrument versus another could affect the data analysis.

The CAS, CDP, and PDI measurements were compared for each IOP when measurements were available (Table S1). The *CAS* measurements were invalid before 6 September 2016 and after 7 October 2018 due to an electronics issue. The CDP-A measurements were invalid for the 2016 and 2017 IOPs due to a misalignment of the optical system. The PDI measurements were invalid for the 2017 and 2018 IOPs due to electrical interference on the aircraft, which affected data transfer between the instrument and onboard computers. Hence, the following sections present analyses comparing measurements from the CAS and the PDI for the 2016 IOP, the CAS and the CDP-B for the 2017 and 2018 IOPs, and the CDP-B and the CDP-C for the 2018 IOP. The measurements collected by the horizontal and vertical channels of the 2D-S, which concurrently sample the cloud volume, were compared for the 2017 and 2018 IOPs.

2016 IOP - CAS versus PDI

Nine research flights between 6 and 27 September 2016 were used to create data sets for comparing measurements from the CAS and the PDI (Table S1). N_c and LWC were calculated for in-cloud measurements, defined as 1 Hz samples with CAS $N_c > 10$ cm^{-3}, PDI $N_c > 10$ cm^{-3}, and King LWC > 0.05 g m^{-3}. The range of the difference between the CAS and PDI data set parameters was defined using the 95 % confidence intervals (CIs) from a two-sample t-test (Table S2 and Table S3). For example, the difference between N_c for in-cloud CAS and PDI data sets was

determined to be between 9 to 12 cm^{-3} with 95 % confidence. The average PDI N_c was 164 ± 90 cm^{-3} and the average CAS N_c was 153 ± 72 cm^{-3}, where the error estimates represent the standard deviation. The PDI N_c and the CAS N_c were well correlated with Pearson's correlation coefficient (R) = 0.88 but their averages had statistically significant differences (Table S2). The PDI more frequently sampled N_c > 300 cm^{-3} and LWC > 0.5 g m^{-3} (1,353 and 3158 1 Hz measurements) compared to the CAS (302 and 25 1 Hz measurements) (Figure S1).

The average PDI LWC was 0.35 ± 0.19 g m^{-3}, and the average CAS LWC was 0.15 ± 0.09 g m^{-3}. The CAS LWC and PDI LWC were well correlated with R = 0.84 but their averages had statistically significant differences (Table S3). The King LWC had an average of 0.28 ± 0.15 g m^{-3} for the in-cloud data. The average PDI LWC was higher than the average King LWC (95 % CIs: 0.06 to 0.07 g m^{-3} higher, R = 0.78) while the average CAS LWC was lower than the average King LWC (95 % CIs: 0.13 to 0.14 g m^{-3} lower, R = 0.80).

Vertical profiles of CAS LWC, PDI LWC, and King LWC were compared against the adiabatic LWC (hereafter LWC_{ad}) (Figure S2) for in-cloud measurements from cloud profiles flown on the six 2016 research flights used for data analysis (Table S1). The average CAS LWC and King LWC were lower than the average LWC_{ad} (95 % CIs: 0.16 to 0.17 g m^{-3} lower for CAS LWC and 0.01 to 0.03 g m^{-3} lower for King LWC). However, the average PDI LWC was higher than the average LWC_{ad} (95 % CIs: 0.04 to 0.06 g m^{-3} higher). The PDI LWC exceeded LWC_{ad} over the entire cloud layer except the top 10 %, the CAS LWC exceeded LWC_{ad} for the bottom 10 %, and the King LWC exceeded LWC_{ad} for the bottom 40 % of the cloud layer. Marine stratocumulus are typically sub-adiabatic due to cloud-top entrainment and droplet evaporation (Gupta et al., 2021) or cloud

water removal by precipitation. Since the LWC_{ad} represents the theoretical maximum for LWC based on the adiabatic model (Section 3), these results suggest the PDI LWC was an overestimate.

The CAS N_c and the PDI N_c had larger differences (with lower R) when the CAS N_c and PDI N_c both exceeded 200 cm^{-3} (Table S2). On the other hand, for about 65 % of the measurements with CAS N_c and PDI N_c < 200 cm^{-3}, the CAS N_c and PDI N_c had insignificant differences while the CAS LWC and PDI LWC had significant differences for the measurements (Table S3). No obvious trends were observed for these differences as a function of altitude or pitch angle (not shown).

The skewness (α) and mean radius (r_1) were calculated for the CAS and PDI in-cloud measurements. r_1 and α were negatively correlated for each probe with R = - 0.59 for the CAS and R = - 0.65 for the PDI (Figure S3). Over 60 % of the samples had α_{PDI} < 2, CAS N_c < 200 cm^{-3}, and PDI N_c < 200 cm^{-3} (Table S2). For these samples, there were insignificant differences between the average CAS N_c and PDI N_c (Table S3), but the PDI LWC was significantly higher than CAS LWC (Table S3). This was because the average PDI r_1 was 2.1 μm higher than the average CAS r_1. The data samples with PDI N_c > 200 cm^{-3} were associated with r_1 < 10 μm and α_{PDI} > 1 (Figure S3). Higher PDI LWC compared to the CAS LWC, King LWC, and LWC_{ad} with statistically significant differences suggests the PDI could be oversampling droplets with $D > r_1$ since LWC is dominated by the contribution of larger droplets. This would explain the statistically significant differences between the CAS LWC and the PDI LWC despite smaller or statistically insignificant differences between the CAS N_c and the PDI N_c. Based on these comparisons, measurements from the CAS were used to characterize droplets with 3 < D < 50 μm for the 2016 IOP in the absence of measurements from the CDP-A.

2017 IOP - CAS versus CDP-B

The CAS and the CDP-B data sets were created using in-cloud measurements defined as 1 Hz samples with CAS N_c > 10 cm^{-3}, CDP-B N_c > 10 cm^{-3}, and King LWC > 0.05 g m^{-3}. For in-cloud measurements collected over 12 research flights during the 2017 IOP, the average CDP-B N_c (192 ± 123 cm^{-3}) and CDP-B LWC (0.18 ± 0.16 g m^{-3}) were greater than the average CAS N_c (181 ± 96 cm^{-3}) and CAS LWC (0.09 ± 0.07 g m^{-3}) (Figure S4). The average King LWC (0.21 ± 0.15 g m^{-3}) was higher than the average CDP-B LWC (95 % CIs: 0.01 to 0.02 g m^{-3} higher, R = 0.68) and the average CAS LWC (95 % CIs: 0.10 to 0.11 g m^{-3} higher, R = 0.78).

For the research flights flown on 30 and 31 August 2017, the average CDP-B N_c (109 ± 39 cm^{-3}) and CDP-B LWC (0.05 ± 0.04 g m^{-3}) were 96 cm^{-3} and 0.16 g m^{-3} lower than the CDP-B N_c and CDP-B LWC averaged over the other flights. The average CAS N_c (146 ± 46 cm^{-3}) and CAS LWC (0.11 ± 0.05 g m^{-3}) for these two flights were 41 cm^{-3} lower and 0.02 g m^{-3} higher than their corresponding averages. The average King LWC for these flights (0.18 ± 0.10 g m^{-3}) was 0.03 g m^{-3} lower than the average King LWC for other flights. Since the relative changes in King LWC and CAS LWC compared to other flights were much smaller, it is unlikely the CDP-B measurements from 30 and 31 August 2017 were accurate. The CDP-B measurements from 30 and 31 August did not impact the results presented in this study since these flights were not included in the data analysis (Table S1) because few cloud profiles were conducted during these flights. However, the data from these flights were excluded from data sets created for comparing the in-cloud CAS and CDP-B measurements for the 2017 IOP.

The 10 research flights between 12 August and 2 September 2017 were used to create data sets for comparing N_c and LWC from the CAS and the CDP-B in-cloud measurements (Figure S4). 95 % CIs between the N_c and LWC from the CAS and the CDP-B are listed in Table S4 and Table S5, respectively. The CDP-B more frequently sampled N_c > 300 cm^{-3} (2536 1 Hz measurements) than the CAS (1623 1 Hz measurements). The average CDP-B N_c was higher than the average CAS N_c with R = 0.91 (Table S4). For 75 % of the samples with *CDP-B* N_c < 300 cm^{-3}, CAS N_c and CDP-B N_c had small differences (95 % CIs: 1 to 5 cm^{-3}) but the average CDP-B LWC and CAS LWC had statistically significant differences (Table S5). This was because the average CDP-B r_1 was higher than the average CAS r_1 (95 % CIs: 1.4 to 1.5 μm higher).

The average King LWC (0.19 ± 0.13 g m^{-3}) was comparable to the average CDP-B LWC (0.18 ± 0.13 g m^{-3}) while the average CAS LWC (0.08 ± 0.06 g m^{-3}) was lower than CDP-B LWC and King LWC. The CAS LWC, CDP-B LWC, and King LWC were compared against LWC_{ad} (Figure S5) for in-cloud measurements from cloud profiles flown on the seven research flights from the 2017 IOP used for data analysis (Table S1). The average LWC_{ad} was greater than each LWC estimate but the differences with CAS LWC (95 % CIs: 0.17 to 0.19 g m^{-3} higher) were higher than with CDP-B LWC (95 % CIs: 0.05 to 0.07 g m^{-3} higher). Thus, measurements from the CDP-B were used to characterize droplets with 3 < D < 50 μm for the 2017 IOP.

2018 IOP - CAS versus CDP-B

For the 2018 IOP, N_c and LWC from the CAS and the CDP-B were compared using data sets created from the in-cloud measurements on six research flights until the CAS was operational (Table S1). These comparisons were consistent with the CAS versus CDP-B comparisons for the

2017 IOP. The average CDP-B N_c (125 ± 92 cm^{-3}) was higher than the average CAS N_c (106 ± 67 cm^{-3}) with statistically significant differences (95 % CIs: 15 to 21 cm^{-3} higher, R = 0.88) (Figure S6). The average CDP-B LWC (0.21 ± 0.14 g m^{-3}) was closer to the average King LWC (0.20 ± 0.12 g m^{-3}) compared to the average CAS LWC (0.10 ± 0.07 g m^{-3}). The average LWC_{ad} was closer to the average CDP-B LWC (95 % CIs: 0.04 to 0.06 g m^{-3} higher) and the average King LWC (95 % CIs: 0.07 to 0.08 g m^{-3} higher) compared to the average CAS LWC (95 % CI: 0.18 to 0.19 g m^{-3} higher). It was hypothesized that the CDP-B provided better estimates of *N(D)* for droplets with 3 < D < 50 μm compared to the CAS for the first six research flights from the 2018 *IOP*.

Based on these comparisons, the CAS could be under-sizing droplets or under-sampling certain droplets during the 2017 and 2018 IOPs. The differences between the data sets from the CAS and the other instruments could be due to droplet co-incidence in the CAS sample volume. It is possible the air flow into the CAS inlet tube could have affected the droplets entering the CAS sample volume compared to the CDP-B sample volume (which had a more open path for droplets). The differences between the estimates of N_c and LWC from the CAS and CDP-B for the 2017 IOP increased slightly when the absolute value of pitch angle exceeded 0.5˚ (Table S4 and Table S5). However, this was not observed for data collected during the 2018 IOP. No obvious trends were observed for these differences as a function of altitude or the skewness from the CAS and the CDP-B *N(D)* (not shown).

2018 IOP - CDP-B versus CDP-C

During the 2016 IOP, cloud probes were installed on newly designed pylons that placed the instruments directly underneath the wing rather than slightly ahead of its leading edge as

commonly regarded as best practice (McFarquhar et al. 2007; Afchine et al. 2018). There was concern that the air flow into a probe sample volume could have been affected by airflow perturbations induced by the wing (Weigel et al. 2016), potentially affecting the size distributions and the calculation of N_c, LWC, and other microphysical parameters. To investigate this, a new pylon was designed at the NASA Wallops Flight Facility and installed on one wing for the 2017 and 2018 IOPs. This pylon placed the CAS and the CDP-B slightly lower and ahead of the leading edge of the aircraft wing, compared to other probes. Therefore, the CDP-B and CDP-C were mounted at different locations relative to the aircraft wing.

The mounting locations of the CDP-B and CDP-C were switched halfway through the 2018 IOP to isolate instrument differences caused by the pylons from those caused by the CDP probes. O'Brien et al. (2021, in prep.) compared the in-cloud measurements from CDP-B and CDP-C and found the mounting position of the probes had only a 6 % impact on the calculation of N_c with the average CDP-B LWC and CDP-C LWC being within 0.02 g m^{-3}. To maintain consistency with the 2017 IOP, in-cloud measurements from the CDP mounted on the new pylon (next to the CAS) were used for data analysis (Table S1) except for 15 October 2018 when the CDP-C, placed on the new pylon, erroneously sampled large N_c due to a qualifier voltage issue. However, the use of measurements from the CDP mounted on the old pylon is unlikely to have a significant impact on the data analysis.

2017 and 2018 IOPs - 2D-S horizontal and vertical channel

N_c and LWC were derived using the in-cloud measurements from the horizontal (N_H and LWC_H) and vertical (N_V and LWC_V) channels of the 2D-S. N_H, N_V, LWC_H, and LWC_V were computed

for 3,966 and 7,612 1 Hz in-cloud measurements with LWC_H and LWC_V between 0.001 to 1 g m^{-3} collected during 7 and 12 research flights from the 2017 and 2018 IOPs, respectively. Based on a linear regression model, N_H and N_V (Figure S7) as well as LWC_H and LWC_V (Figure S8) were highly correlated for the 2017 and 2018 IOPs. Only N_H and LWC_H were available for the 2016 IOP because of soot deposition on the inside of the receive-side mirror of the 2D-S vertical channel. To maintain consistency between the three IOPs, N_H and LWC_H were used in this study despite the availability of N_v and LWC_V for the 2017 and 2018 *IOPs*. The high correlations suggest little difference would have resulted in the data analysis from using the average of the 2D-S channels.

Table S1: P-3 research flights (PRFs) from ORACLES used for data analysis along with instruments that provided valid samples of droplets with $3 < D < 50$ μm during the PRF (primary instrument for data analysis in bold).

PRF date	PRF used	Instruments
Aug 30 2016	No	Aborted flight
Aug 31 2016	No	PDI
Sept 02 2016	No	PDI
Sept 04 2016	No	PDI
Sept 06 2016	Yes	**CAS**, PDI
Sept 08 2016	No	**CAS,** PDI
Sept 10 2016	Yes	**CAS**, PDI
Sept 12 2016	Yes	**CAS**, PDI
Sept 14 2016	Yes	**CAS**, PDI
Sept 18 2016	No	**CAS**, PDI
Sept 20 2016	Yes	**CAS**, PDI
Sept 24 2016	No	**CAS,** PDI
Sept 25 2016	Yes	**CAS**, PDI
Aug 12 2017	Yes	CAS, **CDP-B**
Aug 13 2017	Yes	CAS, **CDP-B**
Aug 15 2017	Yes	CAS, **CDP-B**
Aug 17 2017	Yes	CAS, **CDP-B**
Aug 18 2017	No	CAS, **CDP-B**
Aug 19 2017	No	Aborted flight
Aug 21 2017	Yes	CAS, **CDP-B**
Aug 24 2017	Yes	CAS, **CDP-B**
Aug 26 2017	No	CAS, **CDP-B**
Aug 28 2017	No	CAS, **CDP-B**
Aug 30 2017	No	**CAS**, CDP-B
Aug 31 2017	No	**CAS**, CDP-B
Sept 02 2017	No	CAS, **CDP-B**
Sept 27 2018	Yes	CAS, **CDP-B**, CDP-C
Sept 30 2018	Yes	CAS, **CDP-B**, CDP-C
Oct 02 2018	No	CAS, **CDP-B**, CDP-C
Oct 03 2018	Yes	CAS, **CDP-B**, CDP-C
Oct 05 2018	Yes	CAS, **CDP-B**, CDP-C
Oct 07 2018	Yes	CAS, **CDP-B**, CDP-C
Oct 10 2018	Yes	**CDP-B**, CDP-C
Oct 12 2018	Yes	CDP-B, **CDP-C**

Oct 15 2018	Yes	**CDP-B**, CDP-C
Oct 17 2018	No	CDP-B, **CDP-C**
Oct 19 2018	Yes	CDP-B, **CDP-C**
Oct 21 2018	Yes	CDP-B, **CDP-C**
Oct 23 2018	Yes	CDP-B, **CDP-C**

Table S2: 95 % confidence intervals (CIs) for differences between CAS and PDI N_c (positive when average PDI N_c higher) determined using a two-sample t-test. Number of 1 Hz measurements (n), correlation co-efficient (R) and p-value (p) listed for various criteria applied to the CAS and the CDP $N(D)$, where α refers to skewness. Best-fit slope (M_o) and intercept (C_o) were determined using linear regression for CAS data as a function of PDI data.

Criteria	**n**	**CIs (cm^{-3})**	***R***	**p**	***M_o***	***C_o* (cm^{-3})**
All data	16559	9 to 12	0.88	0	0.70	38
CAS and PDI N_c > 300 cm^{-3}	243	67 to 90	0.46	0	0.12	273
CAS and PDI N_c < 300 cm^{-3}	15147	2 to 5	0.88	0	0.81	24
CAS and PDI N_c > 200 cm^{-3}	4076	32 to 37	0.64	0	0.32	156
CAS and PDI N_c < 200 cm^{-3}	10832	-2 to 1	0.83	0.32	0.82	21
α_{PDI} < 2	14311	4 to 7	0.89	0	0.76	31
α_{PDI} > 2	2248	37 to 48	0.85	0	0.58	60
α_{PDI} < 2, CAS & PDI N_c < 200 cm^{-3}	10066	-3 to 0	0.83	0.06	0.82	21

Table S3: Same as Table S2, but the parameters correspond to comparisons between CAS and PDI LWC. The CIs were positive when the average PDI LWC was higher.

Criteria	**n**	**CIs (g m^{-3})**	***R***	**p**	***M_o***	***C_o* (g m^{-3})**
All data	16559	0.20 to 0.20	0.84	0	0.40	0.01
CAS and PDI N_c > 300 cm^{-3}	243	0.25 to 0.31	0.92	0	0.36	0.02
CAS and PDI N_c < 300 cm^{-3}	15147	0.19 to 0.20	0.84	0	0.42	0.01
CAS and PDI N_c > 200 cm^{-3}	4076	0.23 to 0.25	0.93	0	0.38	0.01
CAS and PDI N_c < 200 cm^{-3}	10832	0.19 to 0.19	0.81	0	0.41	0.01
α_{PDI} < 2	14311	0.21 to 0.21	0.83	0	0.40	0.01
α_{PDI} > 2	2248	0.15 to 0.16	0.92	0	0.35	0.01
α_{PDI} < 2, CAS & PDI N_c < 200 cm^{-3}	10066	0.19 to 0.20	0.79	0	0.41	0.01

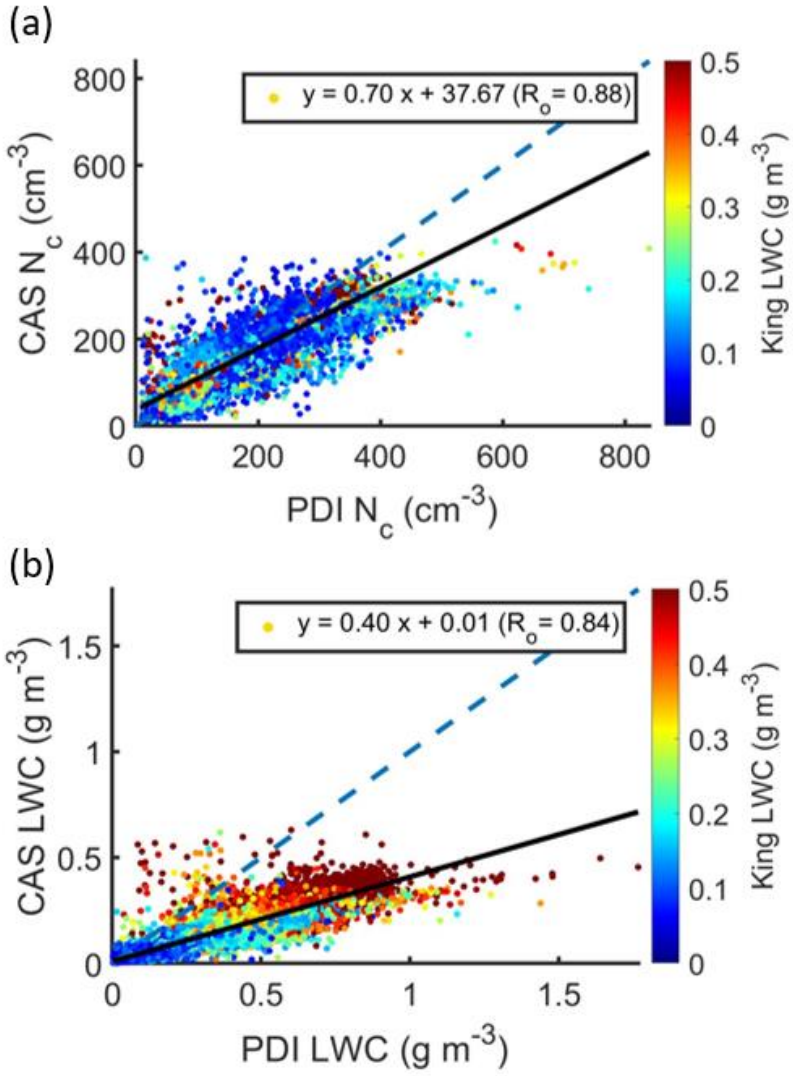
Table S4: Same as Table S2, but the parameters correspond to comparisons between CAS N_c and CDP-B N_c from 10 research flights during 2017 IOP (positive CIs when the average CDP N_c was higher and linear regression coefficients listed for CAS data as function of CDP-B data).

Criteria	n	CIs (cm^{-3})	*R*	p	M_o	C_o (cm^{-3})
All data (excluding 08/30, 31)	11438	16 to 22	0.91	0	0.73	37
CDP N_c > 300 cm^{-3}	2536	73 to 80	0.62	0	0.54	102
CDP N_c < 300 cm^{-3}	8902	1 to 5	0.87	0.01	0.84	20
pitch < - 0.5° or pitch > 0.5°	8445	18 to 25	0.90	0	0.70	42
- 0.5° < pitch < 0.5°	2961	8 to 20	0.91	0	0.80	23

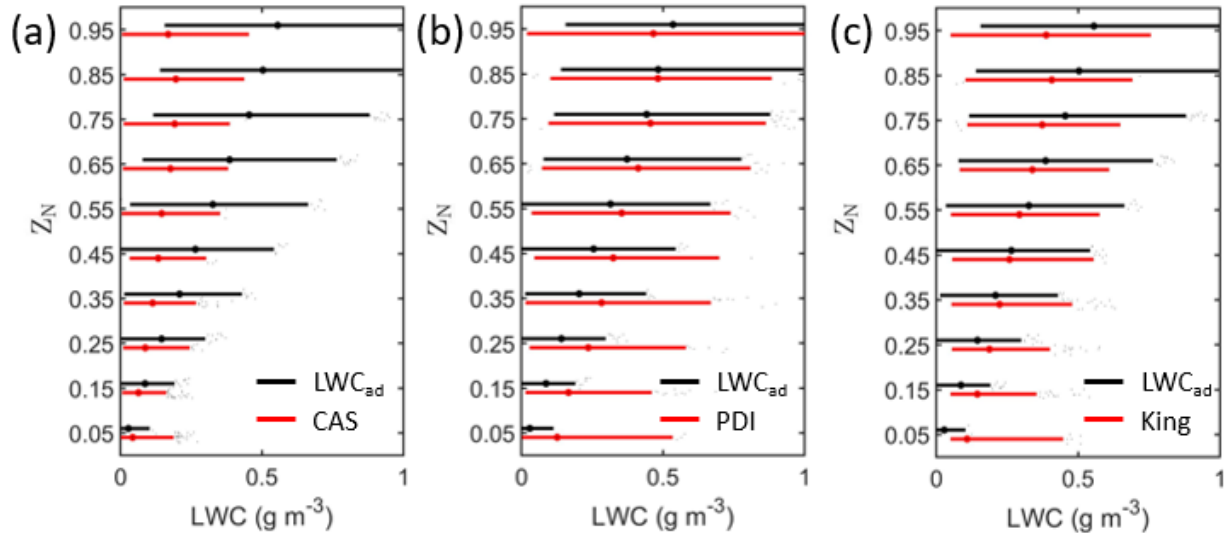
220

Table S5: Same as Table S4, but parameters correspond to comparisons between CAS LWC and CDP-B LWC.

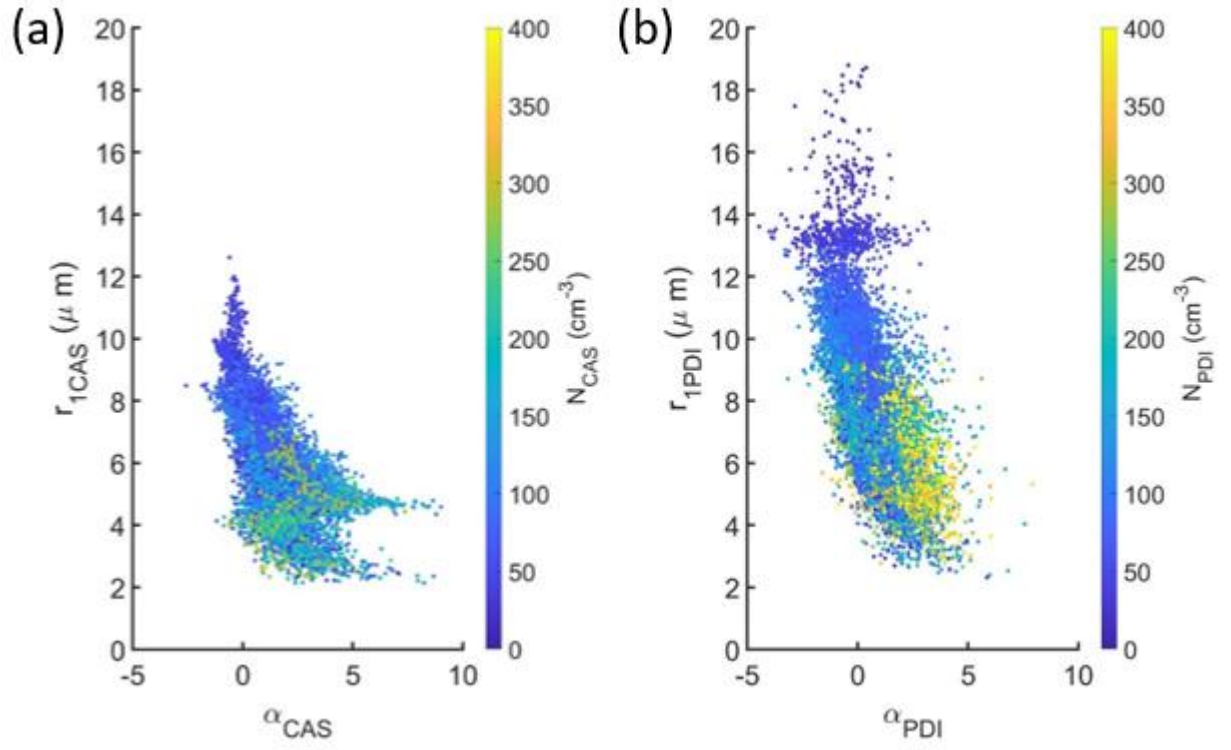
Criteria	n	CI (g m^{-3})	*R*	p	M_o	C_o (g m^{-3})
All data (excluding 08/30, 31)	11438	0.11 to 0.12	0.82	0	0.37	0.01
CDP N_c > 300 cm^{-3}	2536	0.17 to 0.19	0.85	0	0.40	0.00
CDP N_c < 300 cm^{-3}	8902	0.09 to 0.10	0.80	0	0.37	0.02
pitch < - 0.5° or pitch > 0.5°	8445	0.12 to 0.12	0.83	0	0.37	0.01
- 0.5° < pitch < 0.5°	2961	0.10 to 0.11	0.79	0	0.39	0.02



225 Figure S1: (a) N_c and (b) LWC measured by CAS against that measured by PDI during 2016 IOP. Each dot represents a 1 Hz data sample colored by King LWC. Linear regression coefficients indicated in legend.



230 Figure S2: Boxplots representing profiles of (a) CAS LWC, (b) PDI LWC, and (c) King LWC with adiabatic LWC (LWC_{ad}) as function of normalized height above cloud base (Z_N). These data represent cloud samples from cloud profiles flown during the six research flights from 2016 IOP used for data analysis.



235 Figure S3: Mean radius (r_1) versus skewness (α) for (a) CAS and (b) PDI droplet size distributions. Each dot represents a 1 Hz sample colored by the corresponding droplet concentration.

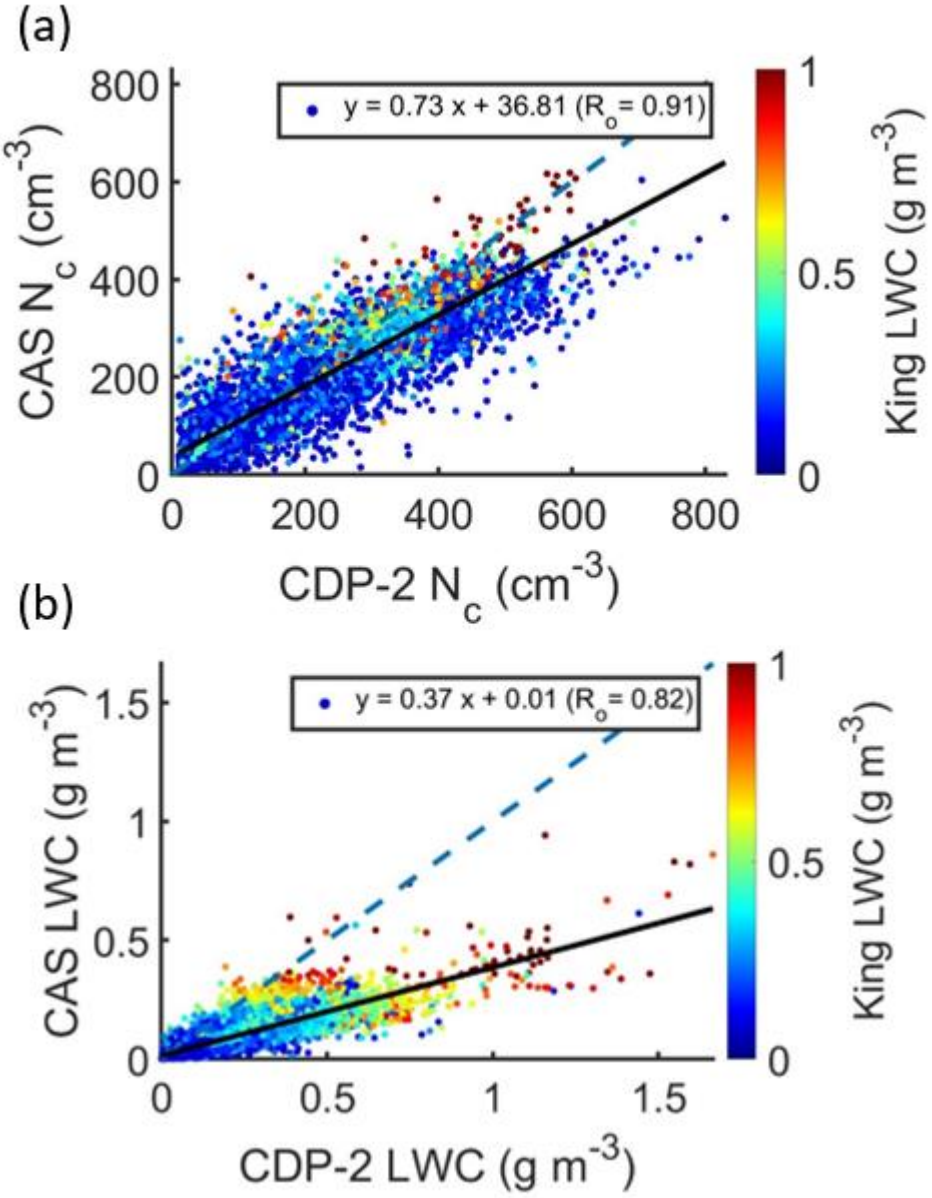


Figure S4: Scatter plots comparing (a) N_c and (b) LWC measured by CAS and CDP-B during 2017 IOP excluding data from 30 and 31 August 2017. Each dot represents a 1 Hz data sample colored by King LWC. Linear regression coefficients indicated in legend.

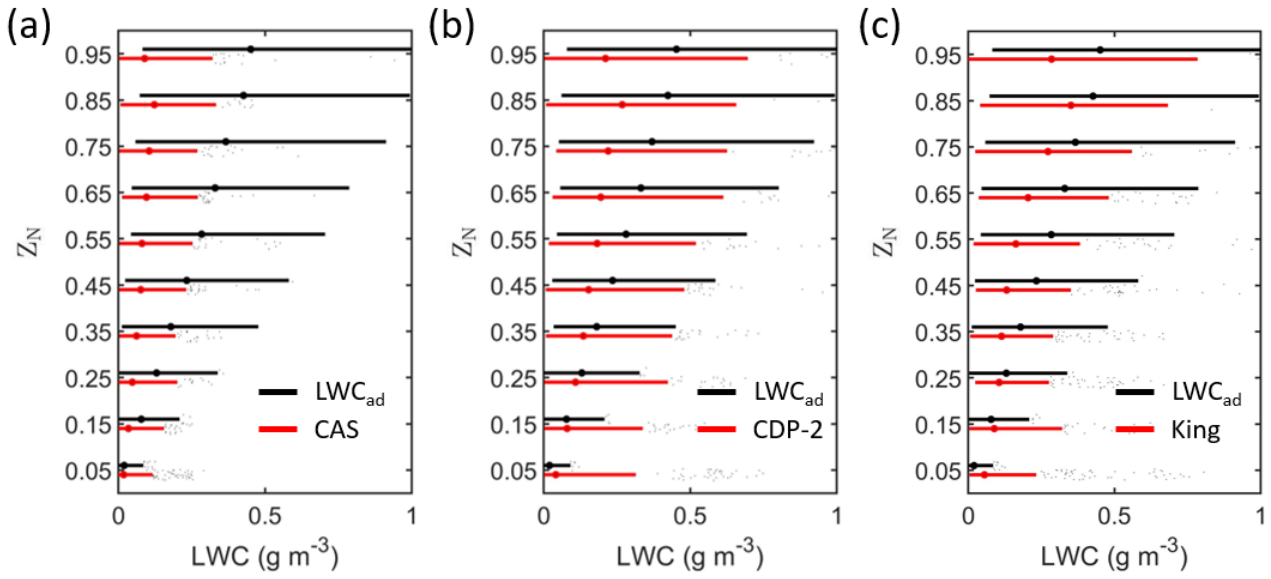


Figure S5: Boxplots representing the vertical profiles of (a) CAS LWC, (b) CDP-B LWC, and (c) King LWC with LWC_{ad} as function of Z_N. These data represent cloud samples from cloud profiles flown during the seven research flights from 2017 IOP used for data analysis.

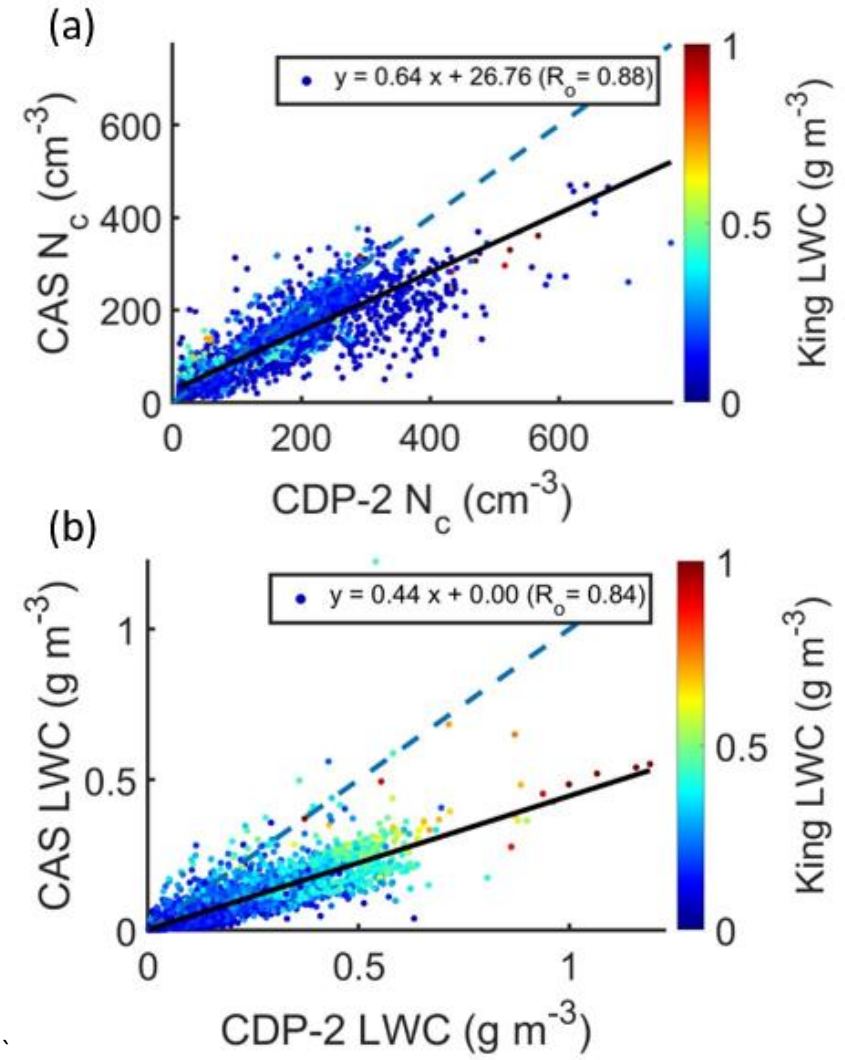
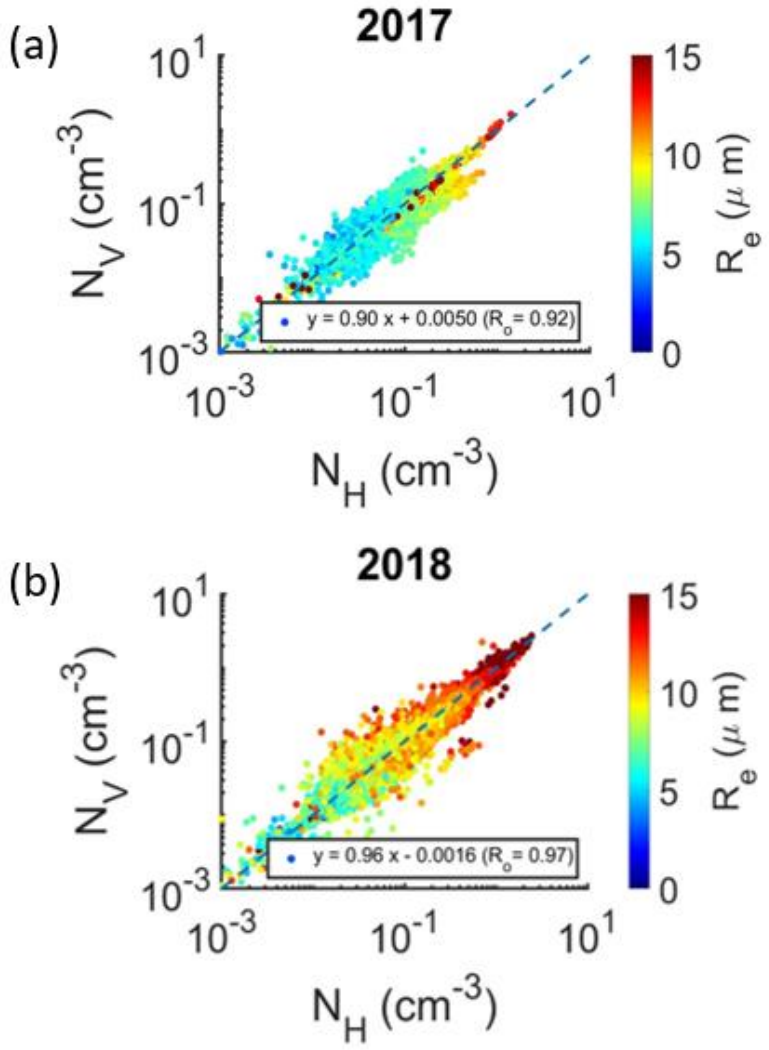


Figure S6: Scatter plots comparing (a) N_c and (b) LWC measured by CAS and CDP-B during 2018 IOP for six research flights when CAS was operational. Each dot represents a 1 Hz data sample colored by King LWC. Linear regression coefficients indicated in legend.



255

Figure S7: Droplet concentration measured by vertical array of 2D-S (N_V) as function of droplet concentration measured by horizontal array of 2D-S (N_H) for (a) 2017 and (b) 2018 IOP. Each data point represents a 1 Hz sample colored by R_e for cloud profiles flown during the research flights from 2017 and 2018 IOP used for data analysis. Linear regression coefficients indicated in
260 legend.

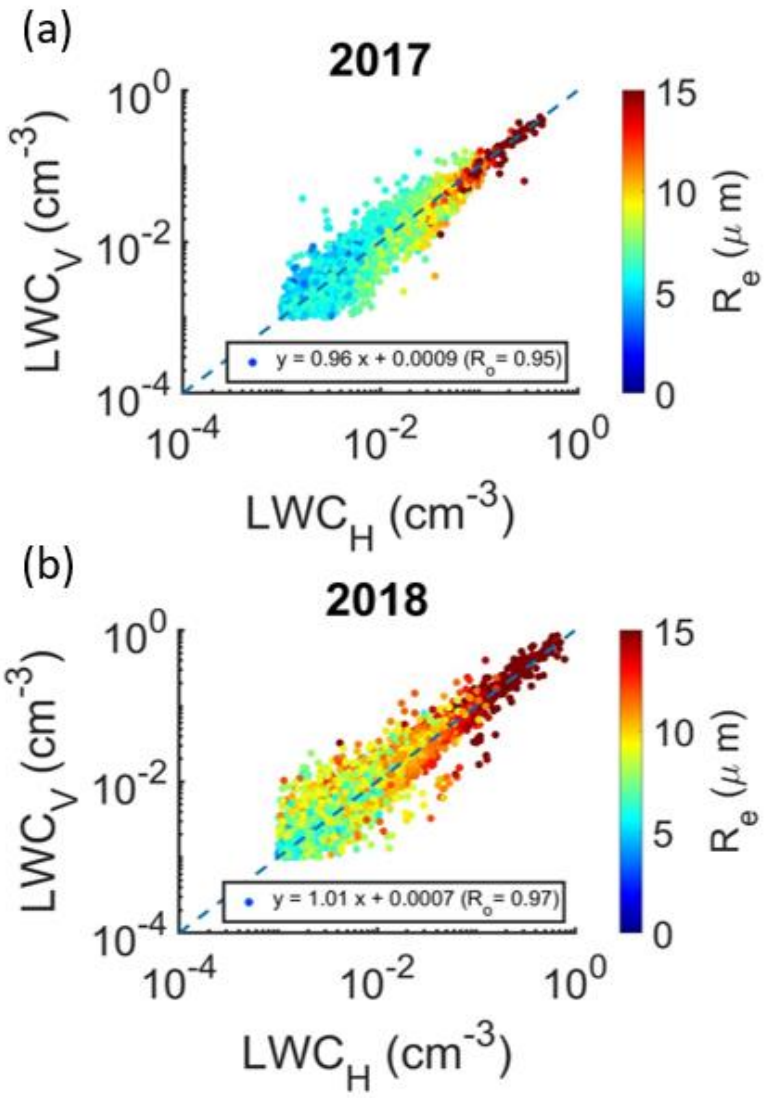


Figure S8: Same as Fig. S7, comparing LWC_H and LWC_V for (a) 2017 and (b) 2018 IOP.